

Il Content Management per la ricerca e la didattica

Abstract

La mole di informazione elettronica afferente alla ricerca e alla didattica che le università e gli enti di ricerca si trovano a gestire è in costante aumento. Nel medesimo tempo, l'evoluzione della tecnologie dell'informazione ha portato alla nascita di nuove soluzioni per l'organizzazione dei contenuti informativi, basate su software che gestiscono la produzione dei contenuti in modo dinamico, flessibile, personalizzabile, ovviando alla rigidità e alla gestione verticistica dei tradizionali siti web. Vengono descritte brevemente l'architettura e le caratteristiche principali di questi software, e vengono presentate alcune possibili applicazioni nel contesto accademico e in quello della ricerca.

Si tratta un'introduzione alle tematiche affrontate dai sistemi di gestione di contenuti ma, soprattutto, di un invito alle biblioteche ad assumere un ruolo primario nella gestione, nella organizzazione, creazione dei contenuti e di servizi informativi per la ricerca e la didattica, impossessandosi delle tecnologie della gestione dei contenuti di ultima generazione e guidandone la scelta e la loro applicazione nei propri enti di appartenenza.

Parole chiave: Sistemi di content management - Gestione di contenuti elettronici per la didattica e la ricerca - Sistemi di document management

The Content Management Systems and their applications in the academic environment

The academic world is affected by a constant growth of electronic scholarly and teaching material requiring permanent or temporary archiving, constant updating, personalized access, frictionless publishing, different output formats, etc. The CMS- Content Management Systems are facing all these problems. The marketplace is offering different solutions; some are very expensive, others are quite inexpensive. This presentation describes the basic architecture and the main features of CMS with specific reference to their applications in universities and research centres. It is important that Italian librarians, in asserting their role as information architects, organizers, familiarize with the different CMS products in order to build their own solutions which better satisfy the information needs of their communities.

Introduzione

In generale quando si parla di *Content Management* e di *Content Management Systems* (d'ora in poi CMS) ci si riferisce alla organizzazione dei contenuti di un sito web complesso, con particolare riferimento al settore del commercio elettronico. Pertanto, con questo nome generalmente si intendono le ultime generazioni di applicazioni sviluppate in casa, o acquisite dal mercato che permettono di gestire siti web, in particolare, le intranet aziendali. In effetti, questa definizione è piuttosto limitativa dal momento che i CMS permettono di gestire in modo integrato molteplici siti web, ma soprattutto permettono di creare, archiviare, gestire, integrare, pubblicare gli stessi contenuti in diverse formati secondo il principio del *write once, publish many*.

Lo scopo di questa relazione è di presentare le caratteristiche principali dei CMS al fine di valutare alcune possibili applicazioni nel contesto accademico e della ricerca. Non rientra nelle finalità di questa relazione discutere dei problemi teorici che sottendono l'applicazione di queste soluzioni in contesti quali gli archivi, biblioteche di conservazione, bensì considerare la loro applicazione nel contesto di una biblioteca dinamica, centrata sull'utente che ha bisogno di informare, comunicare in modo rapido ed efficiente, direi in tempo reale, informazioni integrandole con altri servizi offerti

dalla biblioteca e/o dall'ente e nello stesso tempo assicurandone l'integrità, la sicurezza ed in alcuni casi l'archiviazione permanente.

Cosa sono i CMS?

I CMS non sono un prodotto, bensì un'architettura, un'insieme di attività, di processi relativi alla raccolta, alla gestione, alla pubblicazione di contenuti. Questi sistemi, infatti, si occupano di tutto il processo gestionale dei contenuti, dalla creazione (*authoring*) all'organizzazione del flusso dei contenuti in un'organizzazione lavorativa (*workflow*), alla loro archiviazione (*storage*), infine alla loro pubblicazione (*publishing*). Le funzioni svolte dai CMS si possono riassumere nella:

- o raccolta di contenuti creati o recuperati da fonti esterne);
- o gestione tramite basi di dati/archivi di registrazioni, che contengono componenti di contenuti e o dati amministrativi
- o pubblicazione di contenuti estraendoli da basi di dati al fine di creare pubblicazioni di diversa natura, per esempio un sito web, una pubblicazione a stampa, una newsletter da inviare per posta elettronica, ecc.

In effetti, lo sviluppo del web ed in generale delle tecnologie applicate all'informazione e alla comunicazione ha portato alla nascita di CMS negli ultimi 3/4 anni; si tratta di una tecnologia ancora immatura, nella quale diverse società di software si sono lanciate. Oggi una serie di soluzioni, sono presenti sul mercato, quelle da poche centinaia di euro, quelle da diverse migliaia di euro. Esistono, inoltre, CMS *open source*.

Non esistendo una definizione univoca dei CMS ogni azienda sul mercato dà una descrizione che meglio si adatta alle funzionalità della propria soluzione. In generale le parole chiave utilizzate per descrivere questi sistemi sono: *personalization* (la possibilità di fornire un accesso dinamico personalizzato all'utente, memorizzandone il profilo, le ricerche svolte), *syndication*, (il contenuto del sito può essere arricchito in modo automatico importando materiali da altre fonti, oppure è possibile esportare materiale all'esterno), *asset management* (la gestione di oggetti digitali, testi, suoni, immagini), *re-purposing* (la possibilità di visualizzare gli stessi contenuti in modo diverso, per scopi diversi).

Spesso i CMS vengono confusi con i sistemi di *Knowledge Management* (d'ora in poi KM) e con quelli di *Document Management*, (d'ora in poi DM), di fatto gli esperti del settore distinguono i CMS sia dai primi sia dei secondi, pur riconoscendone delle caratteristiche comuni, mentre nel linguaggio comune spesso vengono usati in modo indifferenziato. In generale i sistemi di KM riguardano la gestione della conoscenza, quindi riguardano prevalentemente la scoperta tramite attività avanzate di *datamining*, di analisi, sintesi di dati, informazioni che permettono di elaborare delle conoscenze laddove i CMS si occupano di raccogliere e di distribuire contenuti. Secondo Bob Boiko, autore del volume "Content Management Bible", potremmo dire che i CMS costituiscono l'infrastruttura nella quale viene raccolta e distribuita la conoscenza portata alla luce, sintetizzata con gli strumenti tipici del KM. Se paragoniamo i CMS ai sistemi di DM potremmo dire che i sistemi di DM costituiscono un sottoinsieme dei CMS, in quanto i sistemi di DM si occupano di gestione di organizzazione e categorizzazione dei files, ne permettono l'archiviazione, l'assegnazione di metadati, il flusso di lavoro, la possibilità di archiviare le diverse versioni del file e di accedere al file tramite lo scorrimento di indici, tramite la ricerca a testo completo. In effetti, questi sistemi andrebbero chiamati File Management System, piuttosto che sistemi di DM. Le funzionalità che li caratterizzano sono presenti anche nei CMS, ma la differenza consiste proprio nel fatto che questi ultimi si occupano non di files, bensì di contenuti, cioè di segmenti di informazione, che vengono prima di tutto creati e poi gestiti. I CMS non si limitano a categorizzare, ad assegnare metadati, perché a differenza dei DMS il loro scopo non è solo quello di organizzare e permettere l'accesso al file/documento bensì di creare informazioni, di pubblicare i contenuti

archiviati, aggregandoli in modo diverso a seconda degli scopi della pubblicazione, del pubblico a cui è destinato.

Architettura e caratteristiche

I contenuti gestiti da questi sistemi si possono definire come informazioni, dati che vengono presentati o pubblicati per essere utilizzati per un certo scopo, e per questa ragione vengono posti in un contesto specifico di presentazione. Questi contenuti richiedono per la loro gestione, la creazione di metadati che descrivano correttamente il contenuto ad essi associato. I contenuti si possono suddividere in argomenti e tipologie; i contenuti si possono suddividere in componenti, pertanto al momento della pubblicazione i vari componenti vengono aggregati con una certa struttura in modo da formare un documento. In effetti, si potrebbe volere ottenere da uno stesso contenuto più pubblicazioni con struttura diversa, pertanto, in questo caso bisognerà dotarsi di una struttura diversa per ciascuna pubblicazione e di una struttura per il contenuto. Si potrebbe inoltre definire che alcuni utenti potranno accedere solo ad alcuni componenti dei contenuti, mentre altri possono accedere a tutti. In generale per ottenere tutto ciò si utilizzano codifiche basate su XML e si suddividono i contenuti in blocchi logici e per quanto attiene le autorizzazioni all'accesso si definiscono sistemi di autenticazione che abilitano o negano l'accesso ad alcuni componenti. I contenuti vengono salvati prevalentemente in una o più basi dati, oppure in file system.

L'architettura di base dei CMS gestisce tutte le funzioni relative a

- la creazione o importazione di contenuti
- l'assegnazione dei metadati
- le attività relative all'accettazione, revisione, approvazione dei contenuti a seconda di come viene organizzato il flusso di lavoro
- la definizione dei parametri di sicurezza per garantire autenticità, sicurezza e integrità agli stessi contenuti onde evitare modifiche non autorizzate, letture non autorizzate, conservazione dell'originale in caso di guasti del sistema o di perdita di dati.
- la integrazione con altre basi dati, con altri sistemi informativi per lo scambio di dati
- la pubblicazione, cioè la possibilità di recuperare parti di contenuti ed altre risorse per creare automaticamente pubblicazioni, per esempio pagine Web, documenti a stampa.

La pubblicazione costituisce un aspetto molto interessante dei CMS. Essa si basa sulla creazione di *templates*, di modelli, cioè di programmi che definiscono come creare la pubblicazione a partire dal contenuto. Questi modelli sono separati dal contenuto, in effetti sono semplici programmi che operano su 3 categorie, la prima relativa agli aspetti statici: testi, alla grafica, a contenuti multimediali, la seconda relativa ai servizi di pubblicazione, cioè al recupero e alla impaginazione di segmenti dei contenuti, personalizzazioni, conversioni di contenuti, costruzioni di indici ecc, la terza riguarda la possibilità di richiamare funzioni presenti su altri sistemi informativi esterni.

Questa caratteristica dei CMS di separare la creazione, il controllo e la responsabilità del contenuto (quello che viene anche chiamato *self service authoring*) dalla sua pubblicazione tramite la definizione di modelli di pubblicazione che possono essere pre-stabiliti, successivamente modificati, costituisce una grossa opportunità per le biblioteche rendendole autonome nella pubblicazione su web. La biblioteca può definire insieme con gli informatici i modelli di cui ha bisogno e poi può procedere in modo autonomo e pubblicare pagine web, newsletter, pubblicazione a stampa senza dover dipendere dall'informatico di turno fino a quando non avrà bisogno di un nuovo modello e allora dovrà ricorrere nuovamente ad un informatico o un bibliotecario con discrete competenze tecniche. Personalmente propendo per la seconda soluzione, auspico la presenza in futuro di bibliotecari *content managers* dotati di notevole autonomia tecnica.

Questa modalità di pubblicazione viene appunto chiamata *frictionless publishing*, in quanto riduce i colli di bottiglia che spesso in diversi posti di lavoro e non solo nelle biblioteche si creano a causa

dell'attesa che intercorre dalla creazione e/o aggiornamento di pagine web o alla loro effettiva pubblicazione su web a causa dell'indisponibilità di tempo del personale tecnico responsabile della pubblicazione delle pagine web.

Le caratteristiche fino ad ora descritte rappresentano le funzionalità di base che ogni CMS generalmente offre, vi sono poi altre funzionalità aggiuntive che costituiscono i punti di forza di un prodotto e ne definiscono meglio la tipologia: sofisticati sistemi di gestione e recupero di documenti, sofisticati sistemi di gestione del versioning, strumenti per la creazione di riviste elettroniche o per la gestione di commercio elettronico, processori XML ecc.

Meritano sicuramente un accenno i metadati amministrativi, descrittivi, tecnici che rappresentano un'attività molto importante e cruciale per il buon funzionamento di un CMS. Una corretta ed efficace descrizione di un contenuto ne permette la sua recuperabilità, un corretto uso, la sua leggibilità da parte di software, la sua possibilità d'uso in futuro. La tendenza spesso ad assegnare metadati di cattiva qualità, per fretta, per mancanza di competenze, e la tendenza a sviluppare i propri metadati in diverse discipline piuttosto che utilizzare standard come il Dublin Core o RDF possono creare seri problemi nel recupero dei materiali e rendono lo scambio di materiali più arduo tra i diversi domini.

Le varie attività svolte dai CMS richiedono l'uso di standard per garantirne l'interoperabilità di pertanto i database, le funzioni di authoring, l'autenticazione degli utenti, le funzioni di riuso/re-purposing dei contenuti si basano tutti gli standard più diffusi.⁽¹⁾

I CMS nelle biblioteche universitarie e di ricerca

L'applicazione di un CMS può essere molto utile all'interno di un progetto di biblioteca ibrida che comprenda i software integrati di biblioteche, l'editoria elettronica acquisita da terze parti o prodotta dall'ateneo/ente di ricerca, le risorse informative bibliografiche, risorse multimediali, materiali didattici. In questo contesto l'architettura dei CMS serve a gestire una parte di questo insieme di risorse e servizi. Certamente un CMS può essere uno strumento molto efficace per la creazione e gestione dei portali delle biblioteche e o dei siti web delle biblioteche, nella gestione di banche dati di materiale multimediale, per migliorare la comunicazione interna tra bibliotecari e bibliotecari ed utenti, per gestire servizi di segreteria degli studenti con i servizi di prestito e di prenotazione offerti tramite l'OPAC della biblioteca. Nella creazione e pubblicazione di bibliografie, manualistica interna per i bibliotecari, per la creazione di materiale didattico per istruire gli utenti all'uso della biblioteca, per la gestione di VRD dinamici. Alcuni CMS possono essere utilizzati anche nella creazione di riviste elettroniche e di archivi istituzionali, quindi possono essere utilizzati se compatibili con Open Archive Initiative per depositare pre-print e/o tesi di dottorato, materiale didattico.

In effetti le funzioni che questi prodotti possono svolgere sono notevoli, ma quello che è necessario fare prima di attuare qualsiasi scelta, è definire quali sono gli obiettivi, quali sono le priorità. Credo che le biblioteche dell'università e della ricerca debbano avere un ruolo leader nell'organizzazione, nell'accesso, nella circolazione, nell'archiviazione e anche nella produzione di contenuti informativi scientifici. Questi contenuti stanno crescendo a dismisura e pertanto richiedono una pianificazione, un controllo, una gestione creano una serie di problemi relativi all'archiviazione permanente, al copyright, al loro recupero, solo per citarne alcuni.. All'estero ci sono interessanti iniziative nelle università olandesi, svedesi, inglesi. Le biblioteche gestiscono al loro interno i centri di produzione digitale degli atenei, questi centri sono considerati sezioni delle biblioteche dell'università., così come nel Regno Unito ci sono biblioteche Library Learning Centres che coordinano le attività della biblioteca e le attività di apprendimento/formazione.

E' necessario, pertanto, definire un progetto, darsi delle linee-guida, creare gruppi di lavoro e/o strutture che si occupino seriamente delle attività di gestione, produzione, di creazione di informazione elettronica. Mi sembra che in Italia, in generale, nell'ambito universitario e della

ricerca si stia procedendo senza alcuna seria pianificazione per quanto afferisce la creazione di contenuti digitali, archivi istituzionali, riviste elettroniche a pagamento o open access, l'uso di piattaforme e-learning, ognuno sperimenta per conto suo. La sperimentazione è sicuramente importante, ma lo è altrettanto la definizione degli scopi ed obiettivi finali.

I CMS sono stati applicati nei Library Learning Centres di alcune università inglesi; il Regno Unito, come ben sappiamo, ha JISC- Joint Information System Committee, la commissione governativa inglese che ha il compito di promuovere le nuove tecnologie dell'informazione e della comunicazione nell'ambito dell'istruzione secondaria ed accademica.

Con riferimento ai CMS, JISC ha commissionato qualche anno fa uno studio nel quale vengono analizzati e valutati le caratteristiche dei CMS, gli sviluppi di questi sistemi, il mercato, ecc.² Lo scopo di questo studio era appunto quello di fornire alle università uno strumento informativo che potesse aiutarle ad operare delle scelte in questo ambito; sebbene lo studio sia stato pubblicato nel settembre del 2001 e da allora nuove soluzioni sono state sviluppate ed immesse sul mercato, esso rimane molto utile ed interessante.

Lo studio mette in evidenza come nella scelta di un CMS è importante avere chiari i propri obiettivi, analizzare le caratteristiche dei prodotti esistenti sul mercato per poter scegliere la soluzione che meglio si adatta al proprio contesto. Viene fatto presente che nessun prodotto soddisfa tutte le esigenze, pertanto, spesso la soluzione migliore è acquisire un prodotto basato su tecnologia standard che assicuri precise funzionalità di base e su questo applicare altri prodotti, sviluppare altre funzionalità, insomma aver l'approccio del *buy and build*.

La relazione intende essere solo un'introduzione alle tematiche affrontate dai CMS, ma soprattutto vuole essere un invito alle biblioteche ad assumere un ruolo primario nella gestione, nella organizzazione, creazione dei contenuti e di servizi informativi per la ricerca e la didattica impossessandosi delle tecnologie di gestione dei contenuti di ultima generazione e guidandone la scelta e la loro applicazione nei propri enti di appartenenza.

1 - In generale gli standard utilizzati sono SQL per i database, WebDAV per l'autoring; LDAP per la gestione degli utenti, CSS, XHTML, XSL, XSLT, RSS per il re-purposing.

2. Browning Paul and Michael Lowndes. *JISC TechWatch Report: Content Management Systems*. September 2001 < http://www.jisc.ac.uk/uploaded_documents/tsw_01-02.pdf >

Bibliografia

Boiko, Bob. *Content Management Bible*. New York, Hungry Minds, 2002, pp.966

Boiko, Bob. *Why content management. A CM domain white paper*. Metatorial Services Inc. & Hungry Minds, 2002

<http://www.metatorial.com/Papers/iqpc_singapore.pdf>

Browning Paul and Michael Lowndes. *JISC TechWatch Report: Content Management Systems*. September 2001

< http://www.jisc.ac.uk/uploaded_documents/tsw_01-02.pdf >

DeBandi, Michele. Sistemi web content management con funzionalità di redazione distribuita. Tesi di laurea. Politecnico di Torino, Settembre 2003

< <http://eprints.rclis.org/archive/00001380/> >

Le nuove sfide dell'e-content management per utenze differenziate. Atti del seminario svoltosi a Roma il 31 ottobre 2003 a Bibliocom dell'AIB a cura di Paola Gargiulo. Roma, AIDA, 2003 pp.63.

Siti Web

CMSWatch. Un sito web che raccoglie informazioni, analisi, rapporti sui sistemi di content management

<<http://www.cmswatch.com>>

La presentazione powerpoint di questa relazione è disponibile sul server di E-lis

<<http://eprints.rclis.org/archive/00001476/>>